Levels of Evidence in the Neurosurgical Literature: More Tribulations Than Trials

Blake A. Yarascavitch, MD*
Jennifer E. Chuback, MD‡
Saleh A. Almenawer, MD*
Kesava Reddy, MD*
Mohit Bhandari, MD, PhD§

*Division of Neurosurgery, ‡Division of Plastic Surgery, and §Division of Orthopedic Surgery, Department of Surgery, McMaster University, Hamilton, Ontario, Canada

Correspondence:
Blake A. Yarascavitch, MD,
293 Wellington St N,
Ste 110, Hamilton,
Ontario, Canada L8L 2X2.
E-mail: blake@yarascavitch.com

Received, January 31, 2012.
Accepted, August 16, 2012.
Published Online, September 14, 2012.

BACKGROUND: The importance of evidence-based medicine has been well documented and supported across various surgical subspecialties.

OBJECTIVE: To quantify the levels of evidence across publications in the neurosurgical literature, to assess the change in evidence over time, and to indicate predictive factors of higher-level evidence.

METHODS: We reviewed the levels of evidence across published clinical studies in 3 neurosurgical journals from 2009 to 2010. Randomized trials were evaluated by use of the Detsky Quality of Reporting Scale. Levels-of-evidence data for the same journals in 1999 were obtained from the literature, and regression analysis was performed to identify predictive factors for higher-level evidence.

RESULTS: Of 660 eligible articles, 14 (2.1%) were Level I, 54 (8.2%) were Level II, 73 (11.1%) were Level III, 287 (43.5%) were Level IV, and 232 (35.2%) were Level V. The number of Level I studies decreased significantly between 1999 and 2010 (3.4% vs 2.1%, respectively; $P = .01$). Seven randomized clinical trials were identified, and 1 trial had significant methodological limitations (Mean Detsky Index = 16.3; SD = 1.8). Publications with larger sample size were significantly associated with higher levels of evidence (Levels I and II; odds ratio, 1.7; 95% confidence interval, 1.45-2.05; $P = .001$). The ratio of higher levels of evidence to lower levels was 0.11.

CONCLUSION: Higher levels of evidence (Levels I and II) represent only 1 in 10 neurosurgical clinical papers in the top neurosurgical journals. Increased awareness of the need for better evidence in the field through education and adoption of the levels of evidence may improve the conduct and publication of prospective studies.

KEY WORDS: EBM, Evidence-based medicine, Neurosurgery

Neurosurgery 71:1131–1138, 2012 DOI: 10.1227/NEU.0b013e318271bc99 www.neurosurgery-online.com

Evidence-based medicine (EBM) combines clinical expertise and judgment, patient preferences and values, clinical circumstances, and the best available research evidence to provide a framework for patient care.[1-3] Over the past 10 years, the term EBM has moved from academia to the front lines of surgery. Several specialties, including plastic surgery, orthopedic surgery, otolaryngology, oral surgery, and neurosurgery, increasingly publish and promote evidence-based guidelines, evidence-based care paths, and evidence-based questions and solutions.[2,4-7]

Recognizing the need, *The Journal of Bone and Joint Surgery* introduced mandatory levels-of-evidence reporting guidelines based on the criteria published by the Oxford Centre for Evidence Based Medicine.[3] These guidelines are included in the instructions for authors and are applied to every clinical paper, assigning a level of evidence (Level I-V) and categorizing studies into therapeutic, prognostic, diagnostic, or economic/decision analysis domains (Table 1). The reliability of the application of levels of evidence among reviewers has been shown to be high.[2] This use of levels-of-evidence reporting by authors and reviewers for publication is increasingly being adopted by surgical journals.

ABBREVIATIONS: EBM, evidence-based medicine; **RCT,** randomized controlled trial

Rothoerl et al[9] published an investigation into the levels of evidence in the neurosurgical literature in 2003. This study assigned a level of evidence to all published clinical papers in 3 major neurosurgical journals for the year 1999. The authors found that < 4% of the literature was Level I evidence.

TABLE 1. Levels of Evidence for Primary Research Question[a]

	Therapeutic Studies	Prognostic Studies	Diagnostic Studies	Economic or Decision Analysis
Level I	High-quality RCT with statistically significant difference or no statistically significant difference but narrow confidence intervals	High-quality prospective study (all patients enrolled at the same point in their disease with $\geq$ 80% follow-up)	Testing of previously developed diagnostic criteria in series of consecutive patients (with universally applied reference "gold standard")	Sensible costs and alternatives; values obtained from many studies; multiway sensitivity analyses
	Systematic review of Level I RCTs (with homogenous study results)	Systematic review of Level I studies	Systematic review of Level I studies	Systematic review of Level I studies
Level II	Lesser-quality RCTs (follow-up $<$ 80%, no blinding, improper randomization)	Retrospective study	Development of diagnostic criteria based on consecutive patients (with universally applied "gold standard")	Sensible costs and alternatives; values obtained from limited studies; multiway sensitivity analyses
	Prospective comparative study	Untreated control subjects from an RCT	Systematic review of Level II studies	Systematic review of Level II studies
	Systematic review of Level II studies or Level I studies with inconsistent results	Lesser-quality prospective study (patients enrolled at different points in their disease/$<$ 80% follow-up		
		Systematic review of Level II studies		
Level III	Case-control study	Case-control study	Study of nonconsecutive patients (without consistently applied reference "gold standard")	Analyses based on limited alternatives and costs; poor estimates
	Retrospective comparative study		Systematic review of Level III studies	Systematic review of Level III studies
	Systematic review of Level III studies			
Level IV	Case series	Case series	Case-control study	No sensitivity analysis
			Poor reference standard	
Level V	Expert opinion	Expert opinion	Expert opinion	Expert opinion
	Cases series without statistical analysis			

[a]RCT, randomized controlled trial. Adapted from *Journal of Bone and Joint Surgery* instructions to authors (http://www.jbjs.org/public/instructionsauthors.aspx).[8] Rating scale used for levels of evidence for primary clinical research question. The level of evidence varies slightly in its definition on the basis of the type of study question being therapeutic, prognostic, diagnostic, or economic and decision-making analysis.

However, since the publication of that article, there has been increased promotion of EBM within neurosurgery. With the support of academic institutions and professional societies, more resources are available than ever before to educate and facilitate evidence-based research and practice within neurosurgery.[10,11]

The use of the levels-of-evidence system as a metric for quality in the neurosurgical literature is largely unknown. Furthermore, little is known about predictors of higher-level evidence. This information is important for quality assurance and as a guiding force for future education and research.

To inform this issue, we conducted a review of the neurosurgical literature over a 1-year period from 2009 to 2010 (hereafter referred to as 2010) to identify the levels of evidence associated with clinical articles. We further aimed to compare current levels of evidence with those from the same journals in 1999 (1 decade earlier) and to examine any predictive factors associated with the publication of articles with higher levels of evidence. We hypothesized that with the increased awareness of the EBM and levels of evidence, the proportions of articles with higher-quality evidence have increased significantly over the past decade.

PATIENTS AND METHODS

Eligibility Criteria

Eligible studies were English language clinical scientific neurosurgery publications (human studies, including case reports, case series, case-control studies, cohort studies, and randomized trials) published in 3 highly cited neurosurgical journals that had published baseline levels-of-evidence distributions. Basic science studies, cadaveric studies, nonsystematic review articles, editorials, correspondence, and expert opinions such as descriptions of novel techniques without clinical examples were excluded. Assessment of eligibility was conducted by 1 reviewer (B.A.Y.).

Search Strategy

All clinical articles published in the area of neurosurgery in the journals *Neurosurgery, Journal of Neurosurgery,* and *Acta Neurochirurgica* from October 2009 through September 2010 were identified through an electronic and manual review of the index from each journal. An approximate 10-year interval was chosen to observe the trends from the baseline report in 1999 to the date of data analysis (September 2010). Our decision to review these 3 journals was based on their high density of general neurosurgical literature with the highest impact factors in the field, providing a good representation of neurosurgical literature currently being published. Additionally, we conducted EMBASE and MEDLINE searches for all randomized clinical trials, controlled clinical trials, meta-analyses, or systematic reviews indexed with the keyword “neurosurgery” in all English language clinical journals indexed for relative comparison of the quantity of Level I neurosurgical evidence being published during this same time period.

Analysis of Articles

Each eligible and included article was evaluated by 2 independent reviewers (J.E.C., B.A.Y.). Reviewers were not blinded to journal, publication date, or authors. Data collected on each article included journal, date of publication, number of authors, geographic region of origin of the primary author, neurosurgical subspecialty, number of centers of collaboration, number of subjects included, study subtype, and level of evidence (Table 1). Articles were categorized by subtype and assigned a level of evidence based on the adapted criteria for authors to *The Journal of Bone and Joint Surgery*.[8] When disagreement occurred between reviewers, the article was discussed and consensus confirmed with a third reviewer (M.B.) with expertise in health research methodology.

We further evaluated the reporting quality of randomized controlled trials (RCTs) using the 21-point Detsky Quality Scale for Randomized Trials.[12,13] This scale is a validated measure for randomized clinical trials that assesses the quality of reporting and methodology of the trial on the basis of randomization, outcome measures, eligibility and exclusion criteria, interventions, and statistical validity of the study.[13] A total of 20 or 21 points (for a positive or negative trial, respectively) are assigned for specific criteria pertaining to the methodology of the trial. The Detsky Quality Scale has been shown to have high interrater reliability and has been used to assess the quality of RCTs from multiple surgical disciplines.[13-15]

Data Analysis

All data were entered into a customized database, and statistical analyses were completed with SPSS 17.0 statistical software (SPSS Inc, Chicago, Illinois). We used descriptive statistics to evaluate levels of evidence across journals. These values were compared with similar data from the same journals in 1999.[9] For the 2010 period, we compared features of higher levels (Levels I and II) and lower levels (Levels III-V) of evidence with χ^2 and Student *t* tests. We also conducted a logistic regression analysis to determine characteristics associated with higher-level studies. All tests were 2 tailed, and we considered $P < .05$ to be the conventional level of statistical significance. To be sufficiently powered (ie, $\beta = 0.20$, $\alpha = 0.05$, 80% study power) to identify a 10% absolute difference in higher-quality studies over time (1999 vs 2010), we required a total of 199 studies in each year.

Comparison With Previous Data

For the purposes of comparison with the prior data of Rothoerl et al[9] from 1999, we recategorized their results on the basis of the current guidelines from *The Journal of Bone and Joint Surgery*. Rothoerl et al had used a very similar grading system for reporting; however, it had been broken down into subgrading of each level of evidence according to study design. For example, Level II evidence was subcategorized into Levels IIa (systematic review of cohort studies), IIb (low-quality RCT or single cohort study), and IIc (“outcome” research). The utility and accuracy of this division are unknown. Therefore, to simplify both comparison and interpretation, we chose to reinterpret their results without this subcategorization.

RESULTS

Literature Search

We identified 1234 articles across the 3 different general neurosurgical journals from October 2009 to September 2010. Approximately half (46.5%) of the articles screened were excluded, the majority of which were correspondence, basic science articles, and literature reviews. Thus, 660 studies met the eligibility criteria and were assessed independently by 2 reviewers. In the assignment of levels of evidence, 8 disagreements occurred that required discussion to achieve consensus. Overall reviewer agreement was excellent ($\kappa = 0.98$; 95% confidence interval, 0.92-0.98). Eligible articles were most commonly published in *Neurosurgery* (40.9%), followed by the *Journal of Neurosurgery* (31.8%) and *Acta Neurochirurgica* (27.3%; Table 2).

Looking for higher-level evidence not captured in the 3 journals represented in our study, we searched EMBASE and MEDLINE for RCTs, controlled clinical studies, meta-analyses, or systematic reviews related to clinical neurosurgery for the year 2010. The search revealed 36 unique articles, of which 16 were relevant and could be classified as Level I (4 studies) or Level II (12 studies) evidence.

Study Characteristics

The majority of eligible studies were conducted in North America (42.0%) and Europe (33.5%; Table 2). Vascular (23.8%) and oncology (23.6%) accounted for half of the eligible published research, followed by general neurosurgery (12.1%), spine (9.1%), peripheral nerve (7.9%), head trauma (6.4%), and functional neurosurgery (5.8%; Table 2). Most articles had 4 or 5 authors (37.7%), but authorship ranged from 1 to 25 contributing authors per article (mean, 5.6; range, 1-25; SD, 2.5). Most studies were therapeutic studies (82.4%), conducted in a single center (90.1%), with ≤ 50 patients (67.1%; Table 2).

Levels of Evidence

Of 660 eligible studies, 14 studies (2.1%) were Level I, 54 (8.2%) were Level II, 73 (11.1%) were Level III, 287 (43.5%) were Level IV, and 232 (35.2%) were Level V (Table 3). The

TABLE 2. Summary of Characteristics From Rated Neurosurgery Studies Published in 2010[a]

	Studies, n (%)		Studies, n (%)
Journal		Study type	
Neurosurgery	270 (40.9)	Therapeutic	544 (82.4)
Journal of Neurosurgery	210 (31.8)	Prognostic	71 (10.8)
Acta Neurochirurgica	180 (27.3)	Diagnostic	42 (6.4)
		Economic and decision analysis	3 (0.5)
Subjects, n		Centers, n	
1	182 (27.6)		
2-10	93 (14.1)	1	596 (90.3)
11-50	168 (25.5)	2	24 (3.6)
51-100	64 (9.7)	3	10 (1.5)
101-200	76 (11.5)	4	6 (0.9)
201-1000	63 (9.5)	5	3 (0.5)
> 1000	14 (2.1)	> 5	21 (3.2)
Subspecialty or discipline		Region	
Vascular/endovascular	157 (23.8)	Africa	2 (0.3)
Oncology	156 (23.6)	Asia	123 (18.6)
General neurosurgery	120 (18.2)	Australia	8 (1.2)
Spine	60 (9.1)	Europe	221 (33.5)
Peripheral nerve	52 (7.9)	India	1 (0.2)
Trauma	42 (6.4)	Middle East	18 (2.7)
Functional	38 (5.8)	North America	277 (42.0)
Other	35 (5.3)	South America	10 (1.5)

[a]Combined characteristics of published clinical articles that were ratable for the 2010 period in *Neurosurgery*, *Journal of Neurosurgery*, and *Acta Neurochirurgica*.

percentages of articles with higher levels of evidence (Levels I and II) published in 2010 were similar in *Neurosurgery* (11.1%) and *Journal of Neurosurgery* (11.4%), with slightly fewer being represented in *Acta Neurochirurgica* (7.8%). Overall, the mean level of evidence across all 3 journals was 4.0 (representing a case series study design), with a mean of 3.9 in the *Journal of Neurosurgery*, followed by 4.0 in *Neurosurgery* and 4.21 in *Acta Neurochirurgica*. The proportion of prospective studies did not differ across journals (Levels I and II; Figure 1).

We identified 7 randomized trials, the majority of which were published in the *Journal of Neurosurgery* (n = 5). Detsky quality scores averaged 80.9% (range, 65%-90.5%). Only 1 trial (14.3%) had major methodological limitations (transformed scores < 75%). The sample sizes across these randomized trials ranged from 17 to 143.

Predictors of Higher Levels of Evidence

Larger sample size studies were significantly more likely to have higher levels of evidence (Levels I and II) than lower ones (odds ratio,1.7; 95% confidence interval, 1.45-2.05; $P = .001$). A prognostic study type was also significantly more likely to be of a higher level of evidence than other study designs (odds ratio, 2.5; 95% confidence interval, 1.80-3.63; $P = .001$). Additionally, multicenter studies were more likely to be of a higher level of evidence (odds ratio, 2.2; 95% confidence interval, 1.12-4.41; $P = .02$).

TABLE 3. Comparison of Neurosurgical Levels of Evidence, 1999 and 2010[a]

	1999,[b] n (%)	2010, n (%)
Reviewer disagreement	9/728	8/660
Clinical studies	346 (35.2)	441 (35.7)
Case reports	287 (29.2)	217 (17.6)
Basic science/experimental	153 (15.6)	171 (13.9)
Technical reports	122 (12.4)	38 (3.1)
Literature reviews	15 (1.5)	56 (4.5)
Other	59 (6.0)	311 (25.2)
Articles screened	982	1234
Included	728	660
Excluded/not ratable	254	574
Level of Evidence		
I	28 (3.8)	14 (2.1)
II	138 (19.0)	54 (8.2)
III	39 (5.4)	73 (11.1)
IV	220 (30.2)	287 (43.5)
V	303 (41.6)	232 (35.2)
Total	728	660

[a]Comparison between 1999 and 2010 time periods across all 3 rated neurosurgical journals demonstrating both the breakdown of the type of clinical study and the distributions of levels of evidence.
[b]The 1999 data were obtained from the published works of Rothoerl et al.[9]

Levels of Evidence Across Specialties

Compared with other medical specialties, the ratio of Level of Evidence I to other levels (Level I/Levels II-V) in neurosurgery was 0.11, which compared favorably with other surgical specialties in the literature, including oral surgery (0.03) and plastic surgery (0.05; Figure 2).[4,6] However, neurosurgery is still lagging behind other

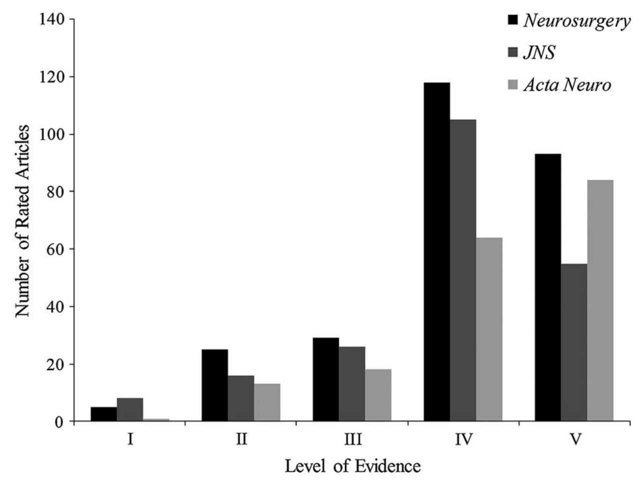

FIGURE 1. *Number of articles for the 2010 period for each journal grouped by level of evidence. Level I articles represented the fewest number of publications from each journal. JNS indicates Journal of Neurosurgery.*

surgical specialties, including orthopedics, ophthalmology, otolaryngology, aesthetic surgery, and urology (Figure 2).[2,5,7,16,18]

DISCUSSION

In this systematic review of neurosurgical articles published in 2010 across 3 prominent neurosurgical journals, Level I studies made up 2.1% of the literature. Factors associated with higher levels of evidence included larger sample sizes, multicenter studies, and a prognostic-type research question.

Strengths and Limitations

Our study represents the most comprehensive and current review of levels of evidence in the neurosurgical literature. Our findings add to previous reviews by further assessing the quality of higher-level evidence in neurosurgery and examining study characteristics associated with higher levels of evidence. Although levels-of-evidence grading was not conducted in a blinded fashion, rating was done independently, and our study has a high degree of interreviewer reliability. We were limited in some aspects of our analysis and comparison with the baseline review of the literature in 1999 by differences in inclusion criteria and levels-of-evidence grading systems. We included the same journals that were assessed in a previous study, which represented the 3 highest-impact-factor neurosurgical journals.[19] For this reason, we were unable to capture neurosurgical studies published in general medical journals or other specialty journals. Finally, because we limited our study to English language publications, our findings may not be generalizable to non-English publications.

Comparison With the Literature

Although the methodology varied slightly between this study and the previous baseline published by Roetherl et al[9] for 1999, some useful comparisons can be made. Important differences include the reporting of distributions of the levels of evidence of the total literature (as opposed to of the included studies); further breakdown of Level I, II, and III studies; and the exclusion of case reports (graded as Level V in our study). To facilitate meaningful comparison and discussion, we have reinterpreted the data from 1999 by the same criteria applied to this study (Table 3).

In 1999, 982 articles were screened between the 3 journals over a 1-year period. By 2010, this had increased to 1234, representing a 25% growth in the number of publications over only a decade. This is reflective of a general increase in the number of medical publications throughout the decade.[20] Similar to 1999, the majority of authors continued to be from North America (Figure 3). This trend is mirrored in other major English language journals.[21,22]

According to this review, the neurosurgical literature has shown a decrease in the proportion of Level I studies over the past decade. Overall, the largest proportion of the literature has shifted from Level V (expert opinion and case reports) in 1999 to Level IV therapeutic or prognostic studies case series in 2010 (Figure 4). Although 3.8% (n = 28) of the literature in 1999 was Level I, this had decreased to 2.1% by 2010 (n = 14). Furthermore, when grouping higher-level (Levels I and II) studies together, we see a decrease from 22.8% in 1999 to only 10.3% in 2010. However, this is still significantly higher than what has been reported in some other surgical specialties, including the general plastic surgical literature over the past 20 years (5.1%) and maxillofacial surgery (2%).[4,6] Specialties that publish a greater proportion of higher-level evidence

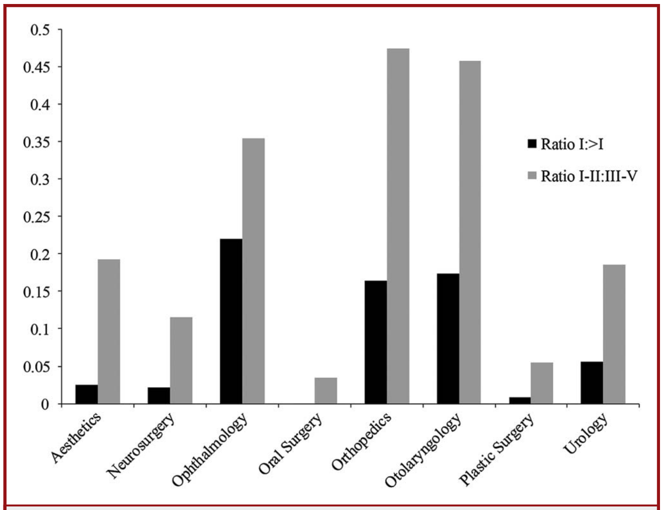

FIGURE 2. *Ratio of publications with higher level of evidence to those with lower level of evidence across surgical specialties. Two thresholds are shown for grouping studies as higher level, a high and low threshold, grouping Level I and Levels I and II, respectively. Data for specialties other than neurosurgery were abstracted from References 18, 5, 6, 2, 7, 4, and 16 (left to right).*

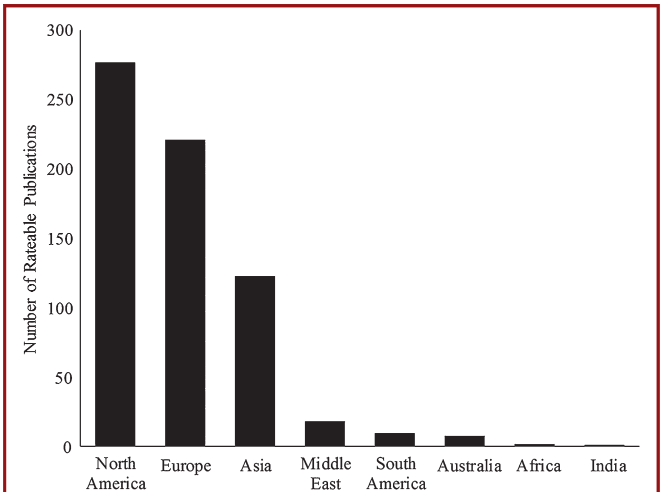

FIGURE 3. *Number of articles for the 2010 period grouped by region of the primary author. The vast majority of articles originated from North America, Europe, and Asia.*

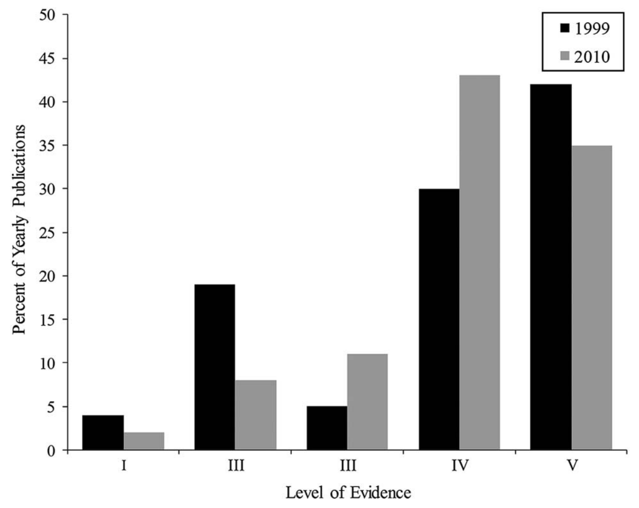

FIGURE 4. *Levels of evidence of publications during the 1999 and 2010 periods as percentages showing relative change in proportion of each level of evidence over a 10-year period.*

include urology (15.6%), ophthalmology (26.2%), orthopedics (32.1%), and otolaryngology (32%).[2,5,7,16]

In an effort to examine the number of publications with higher levels of evidence being published in nonneurosurgical journals, we conducted a literature search to find publications related to neurosurgery that were published in 2010. We only found 16 articles pertinent to neurosurgery that would be considered Level I or II that were not included in our study. This questions the argument that publications with higher levels of evidence concerning the neurosurgical field are consistently being published in nonneurosurgical journals with higher impact factors. The concept of publication bias that selects for submission to highest-impact-factor journals is not new, and it is uncertain whether there have been any changes or trends in this over time.

In our analysis, we found that larger sample size and prognostic studies are predictive of publications with higher levels of evidence, and > 1 center of collaboration was associated with publications with higher levels of evidence. The region of origin, study topic, number of authors, and publishing journal did not have any significant effect on the level of evidence of the publication. This has not been demonstrated previously in the neurosurgical literature. Although our study was adequately powered, the overall very low proportion of articles with higher levels of evidence may be a confounding factor.

Quality of Level I Randomized Trials

Although the number and proportion of Level I studies have decreased over the past decade, the RCTs identified were of high methodological quality. According to our appraisal, only 14% of eligible Level I studies had major limitations. This is an improvement over a 2006-2007 review of 27 intracranial neurosurgical RCTs that found generally suboptimal reporting according to CONSORT (Consolidated Standards of Reporting Trials) guidelines.[23] However, the levels-of-evidence scale used to rate studies in this project automatically downgrades lesser-quality RCTs, which may have led to a selection bias toward methodologically sound RCTs.

Future Directions

We find it interesting that this decrease in higher-level evidence corresponds to a decade of unprecedented interest and exposure to EBM in neurosurgery and across medicine in general.[2-7]

A 2008 editorial published in the *British Journal of Neurosurgery* questioned whether evidence-based neurosurgery is possible.[24] Although acknowledging attitudinal and methodological challenges to producing Level I studies, it recognized the importance of moving past dogma to produce higher-level evidence. Several publications within the neurosurgical literature have addressed the challenges of producing research of higher levels of evidence and have provided guidance to assist with this goal.[10,11,25]

Nonetheless, we are confronted by the reality that despite the availability of these resources, higher-level evidence is decreasing in general neurosurgery journals. From a departmental, societal, and editorial perspective, neurosurgery should concentrate efforts to identify barriers and to promote a movement toward evidence-based practice. Related surgical specialties such as orthopedic surgery, urology, and plastic surgery have faced similar challenges and can serve as role models with proven models for change. In an effort to improve the quality of published literature, *The Journal of Bone and Joint Surgery* adopted the levels of evidence as a standard in its reporting of clinical studies as of 2003. Within only a few years, a review of articles published after the adoption of the levels of evidence reported significant improvements in Level I and II studies.[17] Following suit in 2007, *The British Journal of Urology International* introduced levels-of-evidence grading to its clinical publications. Three years later, a review of the levels of evidence in urology showed that Level I and II studies had doubled from 15.6% in 2007 to 30% in 2010.[16]

After recognizing an area for improvement, an expert panel of society leaders and EBM experts met in the summer of 2010 with the goal of developing a specialty-wide EBM initiative across plastic surgery. In less than a year, a similar level-of-evidence grading system was adopted by *Plastic and Reconstructive Surgery*, the highest-impact-factor journal in the field.[26] At the time of introduction, all 3 journals introduced dedicated EBM sections to engage and educate readership about current best evidence, to foster critical appraisal of the surgical literature, and to disseminate tips about evidence-based surgical practice. To bring about similar transformational change in neurosurgery, we recommend a cross-specialty task force and the adoption of levels-of-evidence grading of publications in neurosurgical journals.

CONCLUSION

Although there has been a shift toward evidence-based surgery over the past decade, there has been a decrease in the proportion of Level I and II studies published across 3 major neurosurgical journals. Level I evidence is still proportionally the least represented. However, these studies are of high methodological quality. Recognizing a need for improvement, a cross-specialty initiative with additional education, emphasis on the conduct of randomized trials and high-quality prospective observational studies, and adoption of a level-of-evidence grading system for the neurosurgical journals is necessary.

Disclosures

Dr Bhandari is funded in part by a Canada Research Chair. The authors have no personal financial or institutional interest in any of the drugs, materials, or devices described in this article.

REFERENCES

1. Dickersin K, Straus SE, Bero LA. Evidence based medicine: increasing, not dictating, choice. *BMJ*. 2007;334(suppl 1):s10.
2. Obremskey WT, Pappas N, Attallah-Wasif E, Tornetta P 3rd, Bhandari M. Level of evidence in orthopaedic journals. *J Bone Joint Surg Am.* 2005;87(12):2632–2638.
3. Center for Evidence Based Medicine. Oxford Centre for Evidence Based Medicine: Levels of Evidence. http://www.cebm.net/index.aspx?o=1025. Accessed July 16, 2010.
4. Loiselle F, Mahabir RC, Harrop AR. Levels of evidence in plastic surgery research over 20 years. *Plast Reconstr Surg.* 2008;121(4):207e-211e.
5. Lai TY, Leung GM, Wong VW, Lam RF, Cheng AC, Lam DS. How evidence-based are publications in clinical ophthalmic journals? *Invest Ophthalmol Vis Sci.* 2006;47(5):1831–1838.
6. Lau SL, Samman N. Levels of evidence and journal impact factor in oral and maxillofacial surgery. *Int J Oral Maxillofac Surg.* 2007;36(1):1–5.
7. Wasserman JM, Wynn R, Bash TS, Rosenfeld RM. Levels of evidence in otolaryngology journals. *Otolaryngol Head Neck Surg.* 2006;134(5):717–723.
8. The Journal of Bone and Joint Surgery. Instructions for Authors. http://www.jbjs.org/public/instructionsauthors.aspx. Accessed July 16, 2010.
9. Rothoerl RD, Klier J, Woertgen C, Brawanski A. Level of evidence and citation index in current neurosurgical publications. *Neurosurg Rev.* 2003;26(4):257–261.
10. Linskey ME. Evidence-based medicine for neurosurgeons: introduction and methodology. *Prog Neurol Surg.* 2006;19:1–53.
11. Bandopadhayay P, Goldschlager T, Rosenfeld JV. The role of evidence-based medicine in neurosurgery. *J Clin Neurosci.* 2008;15(4):373–378.
12. Bhandari M, Richards RR, Sprague S, Schemitsch EH. The quality of reporting of randomized trials in the *Journal of Bone and Joint Surgery* from 1988 through 2000. *J Bone Joint Surg Am.* 2002;84-A(3):388–396.
13. Detsky AS, Naylor CD, O'Rourke K, McGeer AJ, L'Abbé KA. Incorporating variations in the quality of individual randomized trials into meta-analysis. *J Clin Epidemiol.* 1992;45(3):255–265.
14. Thoma A, Farrokhyar F, Bhandari M, Tandan V; Evidence-Based Surgery Working Group. Users' guide to the surgical literature: how to assess a randomized controlled trial in surgery. *Can J Surg.* 2004;47(3):200–208.
15. Colle F, Rannou F, Revel M, Fermanian J, Poiraudeau S. Impact of quality scales on levels of evidence inferred from a systematic review of exercise therapy and low back pain. *Arch Phys Med Rehabil.* 2002;83(12):1745–1752.
16. Borawski KM, Norris RD, Fesperman SF, Vieweg J, Preminger GM, Dahm P. Levels of evidence in the urological literature. *J Urol.* 2007;178(4 pt 1):1429–1433.
17. Hanzlik S, Mahabir RC, Baynosa RC, Khiabani KT. Levels of evidence in research published in *The Journal of Bone and Joint Surgery* (American Volume) over the last thirty years. *J Bone Joint Surg Am.* 2009;91(2):425–428.
18. Chuback JE, Yarascavitch BA, Eaves FF 3rd, Thoma A, Bhandari M. Evidence in the aesthetic surgical literature over the past decade: how far have we come? *Plast Reconstr Surg.* 2012;129(1):126e-134e.
19. ISI Web of Knowledge Journal Citation Reports: Thomson Reuters. http://admin-apps.isiknowledge.com/JCR/JCR. Accessed July 16, 2010.
20. Moorman PW, Schuemie MJ, van der Lei J. An inventory of publications on electronic medical records revisited. *Methods Inf Med*. 2009;48(5):454–458.
21. SCImago, SJR: SCImago Journal & Country Rank. Science analysis. 2007. http://www.scimagojr.com. Accessed August 26, 2011.
22. Holmgren M, Schnitzer SA. Science on the rise in developing countries. *PLoS Biol.* 2004;2(1):E1.
23. Kiehna EN, Starke RM, Pouratian N, Dumont AS. Standards for reporting randomized controlled trials in neurosurgery. *J Neurosurg.* 2011;114(2):280–285.
24. Farah JO, Varma TR. Evidence-based neurosurgery: is it possible? *Br J Neurosurg.* 2008;22(4):461–462.
25. Haines SJ. Evidence-based neurosurgery. *Neurosurgery.* 2003;52(1):36–47; discussion 47.
26. Sullivan D, Chung KC, Eaves FF 3rd, Rohrich RJ. The level of evidence pyramid: indicating levels of evidence in *Plastic and Reconstructive Surgery* articles. *Plast Reconstr Surg.* 2011;128(1):311–314.

COMMENTS

The authors present an interesting review of the current status of the quality of clinical articles published in our major neurosurgery journals as measured by their level of evidence. Their results allow us a moment of self-reflection to think about the big picture of clinical neurosurgery research. Although the number and proportion of randomized controlled trials (RCTs) appear to have decreased since 1999, the news is not entirely bad. First, we cannot definitively conclude from this study that neurosurgeons are doing fewer RCTs. As the authors acknowledge, some of the highest-quality RCTs in neurosurgery are published in general medical journals, not neurosurgical journals. It is possible that a greater proportion of neurosurgery RCTs are being siphoned off by these high-impact general medical journals. Although the authors' preliminary search for such articles revealed only 4 such studies, it is almost certainly the case that a more detailed search would have yielded additional studies. Second, there is some evidence that the quality of RCTs is getting better, although the authors could not compare this directly with the 1999 cohort. Third, the authors' data show us that expert opinion articles no longer make up the largest group of studies, as they did in 1999. The biggest single group is now Level IV studies, and there also appears to be a substantial increase in the proportion of Level III studies. Although this still leaves much to be desired, it shows a trend in the right direction. Aside from what this study documents about changes within neurosurgery, one of the most revealing elements of their article comes from Figure 2. This figure shows neurosurgery to be roughly in the middle of the pack relative to other surgical disciplines but clearly lagging behind some of the leaders in surgery. The Discussion highlights very nicely some of the strategies that other disciplines have used to improve the quality to their literature, some of which could be quite appealing for application in neurosurgery.

Abhaya V. Kulkarni
Toronto, Ontario, Canada

In this very timely article, the authors compellingly demonstrate that in this era of evidence-based medicine, studies with higher levels of evidence (Levels I and II) represent only 1 in 10 neurosurgical clinical articles. The significant drop in the number of Level I studies between 1999 and 2010 is an especially troubling trend. Our medical counterparts such as cardiologists have been doing this successfully for > 20 years. Recently, other surgical specialties such as orthopedics have also made significant changes to increase the number of high-quality prospective studies in their field. This study aptly highlights the question, Why as a specialty are we having trouble organizing, funding, and performing prospective studies? In the last few years, we have started a serious dialog identifying the need for more prospective and comparative effectiveness studies. The present study emphasizes the need to move beyond the dialog and to seriously tackle the barriers that are stopping us from training, funding, and encouraging neurosurgical research in evidence-based medicine.

Chirag G. Patil
Stanford, California